The Folding and Quaternary Structure of Trimeric 2-Keto-3-deoxy-6-phosphogluconic Aldolase at 3.5-Å Resolution†

Irene Moustakali Mavridis‡ and A. Tulinsky*

ABSTRACT: An x-ray crystallographic structure determination has been carried out on 2-keto-3-deoxy-6-phosphogluconic (KDPG) aldolase at 3.5-Å resolution using the multiple isomorphous replacement method with three heavy atom derivatives along with anomalous dispersion contributions from two of the derivatives. Crystals grown from ammonium sulfate–phosphate buffered (pH 3.5) solutions were: cubic, $a = 103.40$ (4) Å, space group $P2_13$. KDPG aldolase consists of trimeric heterologous assemblages utilizing crystallographic threefold symmetry. The overall profile of the oligomeric structure viewed down the threefold axis resembles that of a ship propeller while the subunits are approximate irregular oblate ellipsoids (25 × 45 × 45 Å). The folding of most of the polypeptide chain was traced unambiguously. Secondary structural features consist of nine helical regions (75 residues, 35%) and a pair of two parallel chains. The subunit contains a long empty channel which is about 9 × 9 × 30 Å with one of the pair of parallel chains forming part of the wall. Three mercury binding sites are located in this channel. These might correspond to the two readily accessible and one of the two buried cysteine residues of each subunit. The channel terminates with another cavity of about 8 × 10 × 25 Å near the surface of the oligomeric structure. The regions of the subunits near the threefold axis are characterized by a high degree of secondary structural organization and these make close intersubunit contacts. Quaternary interactions are due mainly to side-chain interactions of helices.

2-Keto-3-deoxy-6-phosphogluconate (KDPG[1]) aldolase, as found in a number of bacteria and isolated from *Pseudomonas putida*, catalyzes the cleavage of KDPG to pyruvate and D-glyceraldehyde 3-phosphate via a class I or Schiff-base assisted mechanism (Meloche and Wood, 1964a; Figure 1). Thus, catalysis proceeds in a manner similar or identical with that effected by fructose-1,6-diphosphate (FDP) aldolase found in higher animals. Heretofore, there has not been a three-dimensional structural description of an enzyme of such function and only two preliminary summaries of x-ray crystallographic studies have appeared, one describing qualitatively the subunit structure of FDP rabbit muscle aldolase (Eagles et al., 1969) and the other describing the subunit arrangement of KDPG aldolase (Vandlen et al., 1973).

Physical (Hammerstedt et al., 1971), chemical (Robertson et al., 1971a), and x-ray crystallographic (Vandlen et al., 1973; Mavridis and Tulinsky, 1975) investigations have shown at KDPG aldolase is composed of three identical subunits and they established the first well-documented case of a trimer oligomer.[2] Since then two additional examples of trimers have been reported: crystalline glucagon (Saski et al., 1975) and a chlorophyll-containing protein from *Chlorobium limicola* (Fenna and Matthews, 1975). As trimeric proteins are so rare in occurrence, a view has developed that there might be fundamental constraints operative against the evolutionary survival of such oligomers as compared with even-numbered ones. This unusual structural aspect of KDPG aldolase has provided additional incentive to the present work.

KDPG aldolase is one of the more thoroughly studied aldolases, structurally (Hammerstedt et al., 1971; Robertson et al., 1971a,b; Barran and Wood, 1971; Mohler et al., 1972; Vandlen et al., 1973) and mechanistically (Grazi et al., 1963; Ingram and Wood, 1966; Rose and O'Connell, 1967; Meloche, 1970; Meloche and Glusker, 1973; Meloche et al., 1975). As a result, it has become a model of aldolytic catalysis comparable to that of FDP aldolase. KDPG aldolase is known to catalyze four chemical reactions (Wood, 1972): (1) cleavage of KDPG; (2) Schiff-base formation between a lysine ϵ-amino group and carbonyl compounds; (3) exchange of solvent protons with methyl hydrogens of pyruvate; and (4) decarboxylation of oxalacetate (Figure 1). The specificity of the enzyme not to form a Schiff base with dihydroxyacetone and hydroxypyruvate, while it does so with hydroxyacetone and α-ketobutyrate, indicates a nonsteric restraint against a hydroxyl group on the deoxy position of KDPG. Apart from this restriction, KDPG aldolase is completely nonspecific in forming Schiff bases with carbonyl compounds. However, Schiff-base formation alone is not sufficient for cleavage. In contrast, the cleavage reaction shows a high degree of specificity requiring the simultaneous presence of a 3-deoxy position, a 4-hydroxyl in the erythro configuration, and a 6-phosphate group.

Crystals and Derivatives

KDPG aldolase was assayed and purified as previously described (Meloche and Wood, 1964b; Robertson et al., 1971a) and crystals of the enzyme suitable for the collection of x-ray diffraction intensity data were grown by the Zeppezauer et al. (1968) method as described by Vandlen et al. (1973).[3] The

† From the Department of Chemistry, Michigan State University, East Lansing, Michigan 48824. *Received April 13, 1976.* This work was supported by the National Science Foundation, Molecular Biology Section (GB-33722 and BMS 75-14171).

‡ Present address: Chemistry Division, Nuclear Research Center, Democritos, Athens, Greece.

[1] Abbreviations used: KDPG, 2-keto-3-deoxy-6-phosphogluconate; FDP, fructose 1,6-diphosphate; tRNA, transfer ribonucleic acid; HgS, mercury(II) succinimide; EHgTS, sodium ethylmercury(II)thiosalicylate; rms, root mean square; ORD, optical rotary dispersion.

[2] Two catalytic trimers associate with three regulatory dimers in aspartate transcarbamoylase (Warren et al., 1973). However, this is a heterooligomeric molecular complex rather than a simple trimeric oligomer.

[3] Carried out in the laboratories of Professor W. A. Wood, Department of Biochemistry, Michigan State University.

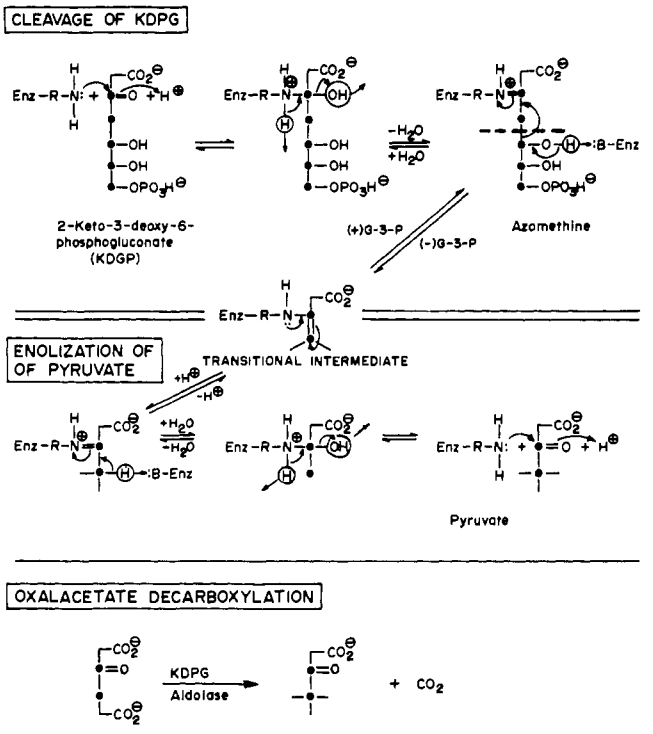
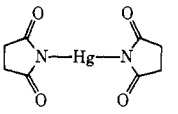
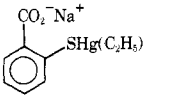


FIGURE 1: Chemical reactions catalyzed by KDPG aldolase.

only change made in the procedure of the latter was to store crystals in 2.5 M $(NH_4)_2SO_4$–0.1 M KH_2PO_4 at pH 3.5 instead of a similar solution but 0.5 M in $(NH_4)_2SO_4$. The crystals were characterized x-ray crystallographically to be: cubic, a = 103.40 (4) Å, space group $P2_13$. From variable equilibrium density measurements and a "salt-free" density (Colman and Matthews, 1971), the molecular weight of protein per unit cell was determined to be 326 000 (Vandlen et al., 1973). Since the space group possesses 12 general equivalent positions, the molecular weight per asymmetric unit is 27 200 ± 2000. This value is in good agreement with other chemical and physical data on the subunit molecular weight. Consideration of the mass per asymmetric unit, the symmetry elements of space group $P2_13$, and the results of chemical and physical solution studies suggested strongly that KDPG aldolase consists of assemblages of four trimers in the crystal with the subunits related to each other by a crystallographic threefold rotation axis (Vandlen et al., 1973). However, 12 monomeric enzyme molecules satisfying the symmetry elements of the crystal could not be excluded as a possibility at that stage of the work. Lastly, the fractional volumes of the unit cell occupied by protein and solvent are about 37 and 63%, respectively, more comparable to those found with tRNA and immunoglobulin crystals rather than to those of globular proteins.

Compounds that produced suitable heavy atom isomorphous substitutions in about 3 weeks of soaking time are listed in Table I along with some of the conditions of the preparation of the derivatives and other data. Two of the compounds, those containing mercury(II), gave derivatives which were stable for periods of up to 3 months so that it was possible to collect 3.5 Å resolution sets of intensity data from these fairly easily. On the other hand, the isomorphous derivative containing $KAuCl_4$ was not very stable, the crystals apparently deteriorating as they soaked in the heavy-atom-containing solution. In addition, these crystals proved to be much more sensitive to radiation damage. Consequently, intensity data collection with this derivative were carried out on only one crystal and only extended to 5.1 Å resolution.

TABLE I: Heavy Atom Isomorphous Derivatives.

Compound	Mole Ratio (Heavy Atom/ Protein)	a (Å)	d_{min} (Å)[a]	Crystals[b]
HgS 	10/1	103.34 (5)	3.5	3
EHgTS 	20/1	103.53 (4)	3.5	2
$KAuCl_4$	10/1	103.44 (7)	5.1	1

[a] Minimum d spacing. [b] Number of crystals used for intensity data collection.

Intensity Data Collection and Processing

The relations among the structure amplitudes for space group $P2_13$ are:

$$|F(hkl)| = |F(\bar{h}kl)| = |F(h\bar{k}l)| = |F(hk\bar{l})| = |F(\bar{h}\bar{k}\bar{l})|$$

and $|F(hkl)| = |F(lhk)| = |F(klh)|$, the cyclic permutation being due to the threefold rotation axis of the cubic system. Thus, only $1/24$ of the diffraction intensities are unique. Since the crystals did not withstand x-ray exposure particularly well and since the number of available crystals was relatively limited, we chose to confine intensity measurements to the unique set of data rather than measuring redundant reflections related by the cyclic permutation and averaging the three measurements. The advantage that could be gained through use of the latter procedure would inevitably be lost in the scaling of the data from a larger number of crystal specimens, whereas a unique set of intensities at 3.5-Å resolution (4577 reflections) could be measured using only two crystals.

The indices of the unique reflections at 3.5-Å resolution were precomputed, sorted in order of azimuthal angle (ϕ) in blocks of increasing Bragg angle (2θ), and stored on a magnetic disc of a computer-controlled diffractometer (Picker FACSI) to be used as identification input for intensity data collection. Data collection in order of increasing 2θ facilitated the combination of data from different crystals and radiation damage to the crystals was corrected for in a natural way: by overlapping the last 100 or so reflections of a previous data collection, the reliability of the measurements and the corrections at the end of latter, where the crystal had suffered the largest amount of radiation damage, could be acertained and confirmed.

Crystals employed for intensity measurements were about 0.5 × 1.0 × 1.0 mm in size and they were mounted in 1.5-mm diameter, thin-walled glass capillaries such that the body diagonal direction of the cubic cell (hhh) was parallel to the azimuthal axis of the diffractometer. Thus, the (hhh) direction was at $\chi = 90°$ and the cubic axes of the crystal were at χ = 35.2° and 120° apart in azimuthal angle.

The intensity data were basically measured using procedures described elsewhere (Vandlen and Tulinsky, 1971; Tulinsky et al., 1973). Some aspects of the data collection of more general interest include the use of a Picker four-circle dif-

fractometer controlled by a Digital Equipment Corp. (DEC) 8K PDP-8 computer (FACS-I system) coupled to a DEC 32K Disc File and an Ampex TMZ seven-track magnetic tape transport. The intensity data were measured by a "wandering" ω-step scan procedure (Wyckoff et al., 1967; Vandlen and Tulinsky, 1971) utilizing balanced Ni/Co filters located between the direct Cu Kα x-ray beam and the crystal to measure background. The alignment of the crystal was monitored periodically every 100 reflections (about every hour) and automatically corrected for with a new orientation matrix if the crystal was considered to be misaligned (10% decrease in intensity of a monitored reflection; 0.04–0.05° in the position of the diffraction bisector; Vandlen and Tulinsky, 1971; Tulinsky et al., 1973). The intensity data collections of the heavy-atom isomorphous derivatives were conducted in similar ways but they were additionally restricted to include *only the observed reflections* of the native enzyme (3853 or 85%). The anomalous differences in the intensities of Bijvoet pair reflections were measured for the HgS and EHgTS derivatives to d_{min} ~5.0 Å. The intensity of a reflection (hkl) was first measured followed by its ($\overline{hkl}$) counterpart. Since Bijvoet differences decreased rapidly with 2θ and were not observable for d_{min} <5.0 Å, the pair measurements were terminated at this point.

The intensity data were also processed using procedures basically described elsewhere (Tulinsky et al., 1973). The following corrections were made to the intensities before they were reduced to structure amplitudes.

(a) Background (Co filter measurement) was empirically corrected for lack of balance primarily due to incoherent scattering.

(b) Net measured intensities were corrected for absorption with an empirical absorption factor which depended on ϕ, χ, and 2θ of each reflection (North et al., 1968). The correction was obtained as a funcion of ϕ from intensity measurements of the (444), (666), (11,11,11), and (16,16,16) reflections at 2θ = 5.9, 8.9, 16.3, 23.9°, respectively. The maximum–minimum absorption ratio of these reflections usually ranged between 1.4 and 1.6.

(c) Net measured intensities were also corrected for decrease due to radiation damage with exposure time. Data collection was usually terminated when the percent decrease in intensity of monitored reflections was about 25–30% ($(I_0 - I_f)/I_0$, where I_0 and I_f are initial and final). However, since the monitored reflections had a 2θ angle in the range 18–24°, most of the reflections did not suffer this amount of radiation damage.

The structure amplitudes of the first native crystal to be measured were used as a reference to which the structure amplitudes of the reflections for the remainder of the crystals were scaled. The scaling was generally carried out in two steps:

(1) All three-dimensional intensity data collections were preceded by a separate 6.0-Å resolution intensity data collection of the centrosymmetric ($0kl$) zone. The 6.0-Å resolution ($0kl$) Patterson projection was used to obtain scaling factors between native crystals and between native and derivative crystals. The ratio of the peak heights of peaks of the Patterson projection of the native crystal, which were not affected by substitution, to the corresponding peaks of the derivatives were estimated and an average ratio was applied to the square of the amplitudes as an initial scale.

(2) The radial intensity distribution of the approximately scaled and averaged derivative amplitudes was then plotted and compared with the native enzyme distribution as a function of $\sin^2\theta$. The derivative amplitudes were fitted with a factor, Q

$$Q = S \exp 2B(\sin\theta/\lambda)^2$$

where S is a multiplicative constant and B is an overall isotropic disorder correction parameter. The scale factors (S) determined in this way for the heavy-atom derivatives did not differ significantly from unity and only the $KAuCl_4$ derivative had an overly large B value (22.5 Å^2), whereas the B values of the HgS and EHgTS derivatives were practically within experimental error (1.0 and 3.5 Å^2, respectively). The problems and difficulties associated with the $KAuCl_4$ derivative which prevented higher order collection (<5.0 Å) have already been mentioned and the large B value is another manifestation of these problems.

A Wilson plot was made corresponding to the logarithm of the ratio of the average $|F|^2$ of the native data to the average sum of the squares of the scattering factors of the atoms of the native enzyme vs. $\sin^2\theta$ in seven ranges.[4] An approximate absolute scale with an overall disorder parameter of about 32 Å^2 was established in this way for the native data. Due to the structure in the $|F|^2$ distribution (maxima at ~9 and 4.5 Å), the Wilson parameters were established using reflections with $2\theta > 19°$.

Finally, the intensities of the ($\overline{hkl}$) reflections of the derivative crystals were processed in exactly the same way as those of the (hkl) reflections. However, the lack of balance and the absorption corrections were measured separately at -2θ angles; in addition, separate distribution curves were used to refine the scaling of the amplitudes to those of the native crystals.

Structure Analysis

(*i*) *Difference Patterson Functions.* Three-dimensional difference Patterson functions of the heavy atom isomorphous derivatives were computed using coefficients of the form $|\Delta F|^2 = (|F|_{PH} - |F|_P)^2$, where $|F|_{PH}$ and $|F|_P$ are the observed structure amplitudes of the heavy atom derivative and the enzyme, respectively. In space group $P2_13$, the twofold screw axes produce vectors at ($\frac{1}{2}$, $\frac{1}{2} \pm 2y$, $\pm 2z$), ($\pm 2x$, $\frac{1}{2}$, $\frac{1}{2} \pm 2z$), and ($\frac{1}{2} \pm 2x$, $\pm 2y$, $\frac{1}{2}$), where (xyz) is the position of a general point. However, due to the threefold rotation axes of the cubic system, the positions and the vectors are also related by a cyclic permutation of coordinates so that the three Harker sections at $u = v = w = \frac{1}{2}$ become equivalent and each contains three Harker peaks relating each coordinate x, y, and z twice. This Harker section of the HgS, EHgTS, and $KAuCl_4$ derivatives is shown in Figure 2 from which it can be seen that: (1) fairly complicated substitutions are occurring in the HgS and EHgTS derivatives; (2) the mercury derivatives have some substitution sites in common; and (3) the $KAuCl_4$ substitution is simple and at a relatively low occupancy.

Several discrete solutions are possible from Harker vectors alone. Selection of a site in one derivative establishes a crystallographic origin to which the coordinates of other sites of the same and the other derivatives can be referred and which also fixes an enantiomorphic configuration. Two prominent sites were found from the Harker section of the HgS derivative (1 and 2 of Figure 2a) and these were also present in the

[4] Two independent sets of native intensity data were collected from first smaller and then larger crystals. The B value between these was 2.0 Å^2 and the residual between the two sets, R, was 0.028, where $R = \sum||F|_1 - |F|_2|/\sum|F|_2$, and $\sum|F|_1 = \sum|F|_2$. Since the data from the larger crystals appeared to be more reliable, the two sets were not averaged and the former were used exclusively in the structure determination.

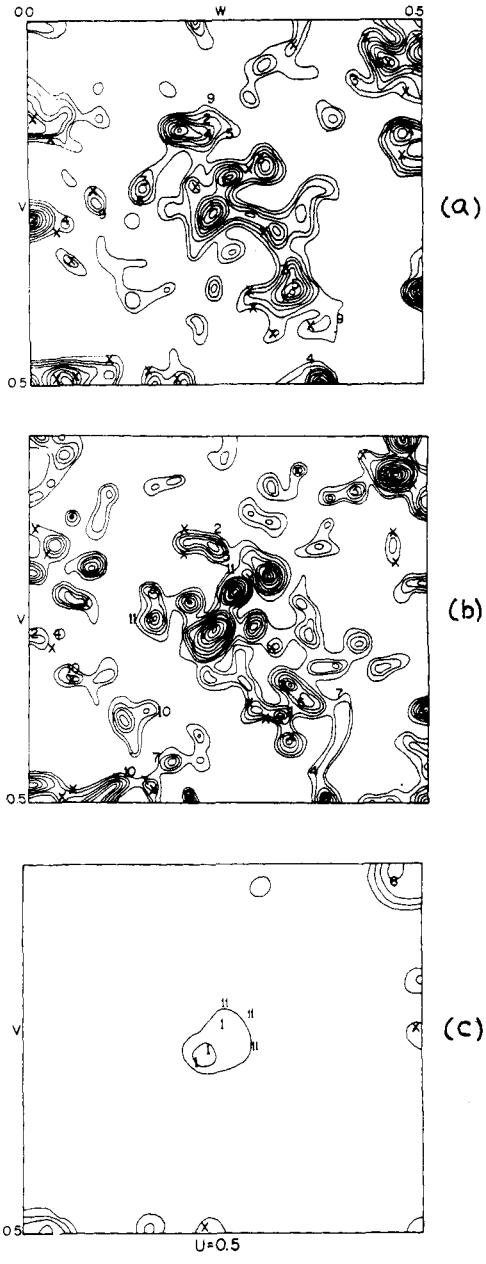


FIGURE 2: Difference Patterson-Harker sections at $u = \frac{1}{2}$ of the three heavy atom isomorphous derivatives. Contours at equal arbitrary intervals above zero; numbers designate self-vector positions; crosses designate cross-vector positions of all sites with principal sites 1, 2, and 3. (a) HgS derivative; (b) EHgTS derivative; (c) $KAuCl_4$ derivative (at half the contour level).

EHgTS derivative (Figure 2b). Moreover, the vectors corresponding to site 2 in the EHgTS derivative were weak. Sixteen possible coordinate combinations of cross-vectors were then examined in order to refer the two sites to the same origin. In addition, it was clear that both derivatives possessed another site of high occupancy. Therefore, site 3 was established with the EHgTS from cross-vectors involving site 1 and then confirmed in the HgS derivative. These results were reconfirmed through the calculation of Kartha's correlation function (Kartha and Parthasarthy, 1965) using coefficients of the form $(|F|_{PH1} - |F|_P)(|F|_{PH2} - |F|_P)$, where 1 and 2 correspond to two different derivatives. All the cross-vectors between sites 1, 2, 3 (HgS) and 1, 3 (EHgTS) were located in Kartha's function. The occupancies of these sites are: HgS (1 has 72, 2 has 62, 3 has 46 electrons) and EHgTS (1 has 70, 2 has 22, 3 has 61 electrons). The remainder of the substitution sites of the two mercury derivatives were easily determined from the foregoing principal sites via the multiple isomorphous replacement method to obtain phase angles used in computing three-dimensional electron and difference electron density distributions of the derivatives. A difference electron density distribution of the $KAuCl_4$ derivative showed two prominent peaks of about equal occupancy at positions similar to sites 1 and 6 of the mercury derivatives. The positions also proved to be consistent with the difference Patterson map of the derivative (Figure 2c).

TABLE II: Heavy Atom Parameters of the Isomorphous Derivatives.

Compound	Site	Position in Fractional Coordinates x	y	z	Occupancy in Electrons	B (Å^2)
HgS	1	−0.1204	−0.1453	−0.8743	72.1	32.5[a]
	2	−0.0022	−0.1798	−0.8872	62.0	29.5[a]
	3	−0.6502	−0.9228	−0.6631	46.0	35.7[a]
	4	−0.0268	−0.1800	−0.1220	39.3	34.1[a]
	5	−0.1719	−0.1235	−0.6570	15.4	37.0
	6	−0.7760	−0.7760	−0.7760	10.7	37.0
	9	−0.0457	−0.2001	−0.8807	9.8	37.0
EHgTS	1	−0.1208	−0.1432	−0.8727	69.6	21.4[a]
	2	0.0016	−0.1726	−0.8802	22.3	21.4[a]
	3	−0.6472	−0.9204	−0.6628	60.7	20.9[a]
	4	−0.0299	−0.1797	−0.1197	23.4	24.1[a]
	5	−0.1712	−0.1136	−0.0727	5.0	30.0
	6	−0.7778	−0.7778	−0.7778	24.9	39.4[a]
	7	−0.1953	−0.0255	−0.5764	8.5	30.0
	10	−0.0898	−0.4746	−0.3127	10.3	30.0
	11	−0.1651	−0.1244	−0.9367	7.7	30.0
$KAuCl_4$	1	−0.1133	−0.1428	−0.8704	30.2	45.0
	6	−0.7667	−0.7667	−0.7667	24.2	45.0
	11	−0.3768	−0.1489	−0.8395	5.5	45.0

[a] Varied 2 cycles during refinement.

(*ii*) *Phase Refinement.* The multiple isomorphous replacement phase angles were improved upon by refining the occupancies and the coordinates of all the heavy atoms by the method of least-squares.[5] The coordinates of the substitution sites did not change much with refinement (~0.2 Å on the average) whereas the occupancies of the more principal sites changed considerably. Refinement of occupancies and coordinates of all the heavy atom sites resulted in an overall improvement and gave a value of the average figure of merit ($\langle m \rangle$) of 0.70 (Blow and Crick, 1959); furthermore, 3016 of 3853 (78%) reflections had an m greater than 0.5 ($N(0.5)$). Isotropic temperature factors of the major sites of the mercury derivatives were varied for two cycles of refinement. The changes were relatively small and converged by the second cycle and improved $\langle m \rangle$ to 0.71 and $N(0.5)$ to 3025. This refinement was followed by three cycles of least-squares refinement of occupancies and coordinates which showed that refinement had converged with respect to these parameters ($\langle m \rangle = 0.71$ and $N(0.5) = 3038$). Table II lists the final parameters of all the heavy atom sites used for the protein phase determination. The low limit in the occupancy determination of a heavy atom site was about five electrons but the error in the determination of higher occupancies was considerably less.

[5] The phase determination and refinement program was space group general and originally written by Professor M. G. Rossmann, Department of Biological Sciences, Purdue University. We would like to thank Dr. Rossmann for supplying us with a copy of the program.

TABLE III: Summary Statistics of Least-Squares Refinement of Heavy-Atom Parameters.[a]

d_{min}	20 Å	10 Å	6.7 Å	5 Å	4 Å	3.3 Å	Overall
			(1) HgS Derivative				
κ	0.14	0.11	0.09				
rms closure E	80	66	60	50	64	64	
rms $\|F_H\|$	506	397	313	266	223	183	
rms $\|\Delta F\|_o$	431	296	216	160	158	131	
AD residual	87	90	92				
$R_{modulus}$	0.155	0.146	0.160	0.147	0.220	0.287	0.197
R_{wgtd}	0.022	0.024	0.031	0.030	0.070	0.104	0.049
R_Δ	0.183	0.186	0.220	0.245	0.313	0.404	0.256
Ref/zone	26	193	473	887	1333	889	
			(2) EHgTS Derivative				
κ	0.11	0.08	0.06				
rms closure E	102	108	85	64	82	78	
rms $\|F_H\|$	448	362	280	248	227	195	
rms $\|\Delta F\|_o$	354	284	200	165	172	157	
AD residual	92	93	94				
$R_{modulus}$	0.219	0.259	0.269	0.217	0.304	0.335	0.279
R_{wgtd}	0.044	0.077	0.079	0.058	0.113	0.139	0.092
R_Δ	0.278	0.323	0.368	0.325	0.405	0.416	0.374
Ref/zone	26	193	467	858	1321	825	
			(3) $KAuCl_4$				
rms closure E		71	65	75			
rms $\|F_H\|$		112	93	74			
rms $\|\Delta F\|_o$		127	92	85			
$R_{modulus}$		0.617	0.644	0.918			0.774
R_{wgtd}		0.388	0.474	0.997			0.669
R_Δ		0.547	0.650	0.794			0.708
Ref/zone		193	467	822			

[a] N = number of reflections; $|F_P|$ = native protein structure amplitude; $|F_{PH}|$ = heavy atom derivative structure amplitude; $|F_H|$ = calculated heavy atom structure amplitude; $|\Delta+|_o$ and $|\Delta+|_c$ = observed and calculated Bijvoet differences, $||F(+)| - |F(-)||$; $\kappa = |\Delta\pm|_o|F_{PH}|/2(|F_P||F_H|)\sin(\alpha_P - \alpha_H)$; rms closure $E = \sqrt{\Sigma(|F_{PH}| - |\vec{F}_P + \vec{F}_H|)^2/N}$; rms $|F_H| = \sqrt{|F_H|^2/N}$; rms $|\Delta F|_o = \sqrt{\Sigma(|F_{PH}| - |F_P|)^2/N}$; AD residual = $\sqrt{(\Sigma(|\Delta\pm|_o - |\Delta\pm|_c)^2/\Sigma|\Delta\pm|_c^2)}$; $R_{modulus} = \Sigma||F_{PH}| - |\vec{F}_P + \vec{F}_H||/\Sigma|F_H|$; $R_{wgtd} = \Sigma\omega(|F_{PH}| - |\vec{F}_P + \vec{F}_H|)^2/\Sigma\omega|F_H|^2$, $\omega = 1/E^2$; $R_\Delta = \Sigma||F_{PH}| - |\vec{F}_P + \vec{F}_H||/\Sigma||F_{PH}| - |F_P||$.

Similar positional sites in different derivatives are denoted by the same site number in Table II.

(iii) Determination of the Absolute Configuration. A right-handed system of coordinates had been used for indexing the reflections during intensity data collection. The interpretation of the difference Patterson function is consistent with either of two centrosymmetrically related sets of heavy atom coordinates but only one of these is consistent with the indexing system adopted. In the present case, the ambiguity was removed by including the anomalous dispersion contributions of HgS and EHgTS in structure factor calculations and comparing the latter with the observed Bijvoet differences (Lipscomb et al., 1966). Initially, $\langle m \rangle$ was 0.715 for all reflections and $N(0.5) = 3063$. Upon interchanging the roles of $|F_{PH}(hkl)|$ and $|F_{PH}(\overline{hkl})|$ in the calculation of the difference anomalous dispersion contribution, $\langle m \rangle$ increased to 0.720 and $N(0.5)$ to 3082 (80%), suggesting that the latter coordinates correspond to the correct absolute configuration. Since the anomalous dispersion measurements extend only to about 5-Å resolution, $\langle m \rangle$ for all reflections is not a fair measure of the improvement effected. Upon examination of regions with minimum spacing of 20, 10, and 6.7 Å containing 27, 95, and 476 reflections, respectively, the first configuration gave $\langle m \rangle$ = 0.881, 0.856, and 0.801, while the second configuration gave 0.909, 0.875, and 0.832 for each of the foregoing ranges. Therefore, the most probable assumption is that the centrosymmetric coordinates of those originally derived give the correct absolute configuration. These coordinates are listed in Table II. The effect of the choice of the correct absolute configuration on $\langle m \rangle$ is shown in Figure 3 along with the distribution of the reflections with respect to the figure of merit and the difference in this distribution upon interchanging the handedness of the heavy atom coordinates.

A summary analysis of the least-squares refinement after anomalous dispersion was applied with the correct hand is given in Table III. The κ ratio, which is an experimental estimate of $\Delta f''/(f + \Delta f')$, where $\Delta f'$ and $\Delta f''$ are the real and imaginary corrections of the anomalous scattering components of the scattering factor, f, of the heavy atoms, is in close agreement to the expected theoretical value of 0.12. The mean scattering contribution of the heavy atoms (rms $|F|_H$) is generally considerably greater than the rms closure error E, except for the last scattering range of the $KAuCl_4$ derivative. The R agreement indices of the mercury derivatives are probably as favorable as can be expected, while those of the $KAuCl_4$ derivative are not, once again reflecting the difficulties mentioned earlier associated with this derivative.

Results and Discussion

(i) Electron Density Maps. "Best" electron density maps including reflections with $m > 0.3$ were calculated at 3.5-Å resolution in planes perpendicular to the a axis and in planes perpendicular to a threefold axis of the crystal. A value of $F(000)/V = 0.127e$ Å^{-3} was also used in the calculations based on the amino acid composition of KDPG aldolase. The rms error of the "best" electron density (Blow and Crick, 1959) was calculated to be $0.088e$ Å^{-3}. Contours of the electron density were drawn beginning at $0.3e$ Å^{-3} ($>3\sigma(\rho)$). The

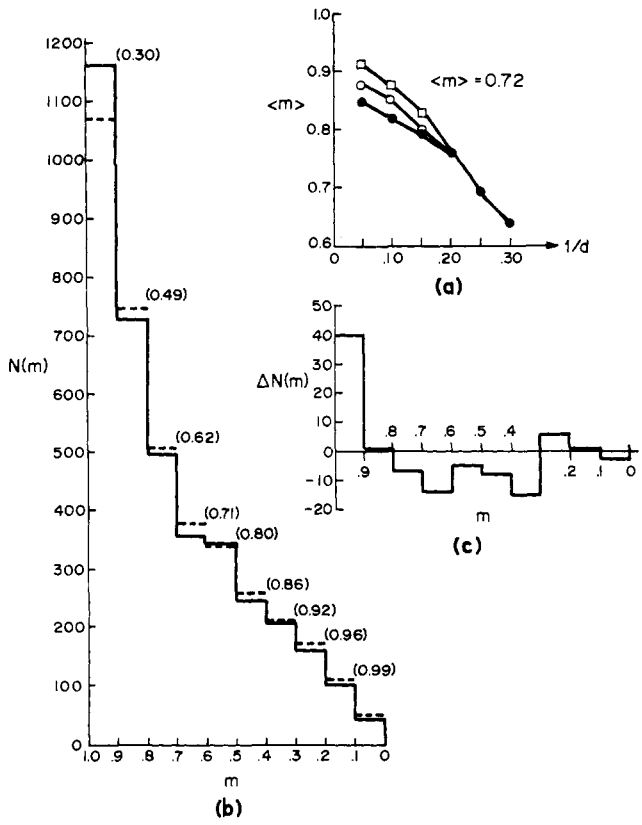


FIGURE 3: Summary statistics of phase-enantiomorph determination. (a) Average m with resolution; (●) without anomalous dispersion (AD); (○) with AD, incorrect hand; (□) with AD, correct hand. (b) Distribution of m among reflections; (solid line) with AD correct hand; (broken line) without AD; fraction of reflections with m greater than given m in parentheses. (c) Difference in number of reflections with m on interchanging hand of heavy atom coordinates.

largest value of the electron density was about $0.74e$ Å^{-3} ($\sim 8.5\sigma(\rho)$).

The electron density maps proved to be consistent with the deduction based on the variable equilibrium density measurements (Vandlen et al., 1973) which suggested that a large fractional volume of the unit cell (63%) was occupied by mother liquor. The maps contained large regions with a mean electron density less than $0.3e$ Å^{-3}, while regions of electron density greater than $0.4e$ Å^{-3} generally formed well-defined and clearly connected peaks. The latter have been interpreted in terms of the protein molecules in the crystal.

(*ii*) *Molecular Packing.* Each protein subunit possesses close intermolecular contacts with four others thus forming a three-dimensional network which extends throughout the crystal. Therefore, crystalline KDPG aldolase is *not an arrangement of monomeric molecules* satisfying $P2_13$ symmetry but rather, an arrangement of trimeric assemblages with a close contact with other trimers. Moreover, it proved fairly easy to distinguish the monomeric protein subunits of the enzyme. These appear to be irregular ellipsoids of approximate dimensions 25 × 45 × 45 Å. The arrangement of the monomeric protein molecules which constitute the asymmetric units in the unit cell is shown schematically in Figure 4. They are related to each other by threefold rotation axes and three mutually perpendicular nonintersecting twofold screw axes. Two different trimeric arrangements of monomers around each threefold axis make relatively close contacts. Three subunits (AAA, BBB, . . .) are situated around each of the four threefold rotation axes to give four trimeric assemblages (Figure 4).

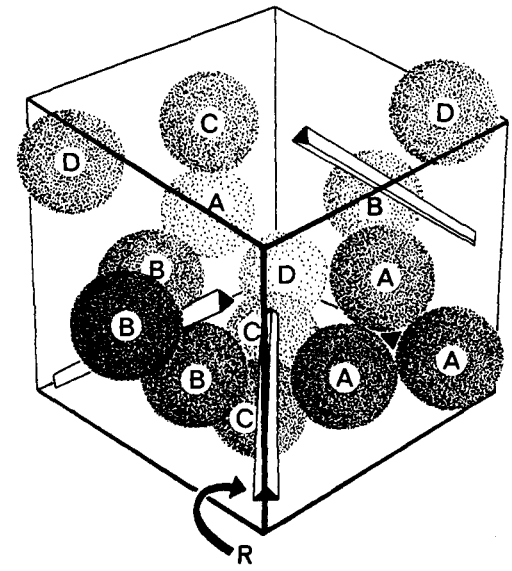


FIGURE 4: Arrangement of asymmetric units in the unit cell of KDPG aldolase. Trimers of first kind (AAA, BBB, . . .); trimers of the second kind (ABC, BCD, . . .); threefold axis (R) is body diagonal; other threefold axes pass through cell edges and cell faces; trimers of both kinds are related to each other by three mutually perpendicular nonintersecting twofold screw axes.

Each subunit of the foregoing additionally interacts with two other subunits, each from a different trimeric assemblage of the first kind, to give a second kind of trimeric arrangement (ACD, BCD, . . .) around the threefold axes.

Since each subunit of trimers of the first kind makes additional close contacts with two other subunits, an ambiguity arises in the choice of the trimeric molecule of KDPG aldolase. The two kinds of trimer differ in the way different portions of their surface interact. The trimers of the second kind are arranged tetrahedrally near the corners of the unit cell (Figure 4). The crystallographic relation between the different trimers is such that the subunits of the first kind of trimer are related to the subunits of the second kind by the three different nonintersecting mutually perpendicular twofold screw axes. In contrast, trimers of both the first and second kind are related to themselves by only one of the three twofold screw axes.

It is impossible to resolve the trimeric ambiguity from symmetry considerations alone. However, a study of the contacts and the interactions of the subunits in each of the two kinds of trimer biases the ambiguity in a definitive way. The general indication of the electron-density map is that the trimer of the first kind is probably the trimeric assemblage in solution (Figure 5): it shows many more and closer contact interactions between the subunits than trimers of the second kind and a relatively small number of intertrimer contacts. The latter thus become a trimeric assemblage of trimers around a threefold rotation axis. Thus, KDPG aldolase is another example of the simplest form of heterologous subunit association (Monod et al., 1965; Hanson, 1966), similar to that observed with glucagon (Saski et al., 1975) and bacteriochlorophyll protein (Fenna and Matthews, 1975).

(*iii*) *Folding and Shape of the Molecule.* Although the electron density map indicates some close intersubunit and intermolecular contacts (Figure 5), nevertheless, it was possible to discern the outline the subunits and their arrangement, both of which can be seen from Figure 6. The overall profile of the oligomeric structure viewed down the threefold rotation axis resembles that of a ship propeller (Figures 5 and 6), while the outline of the subunit approximates an irregular oblate ellip-

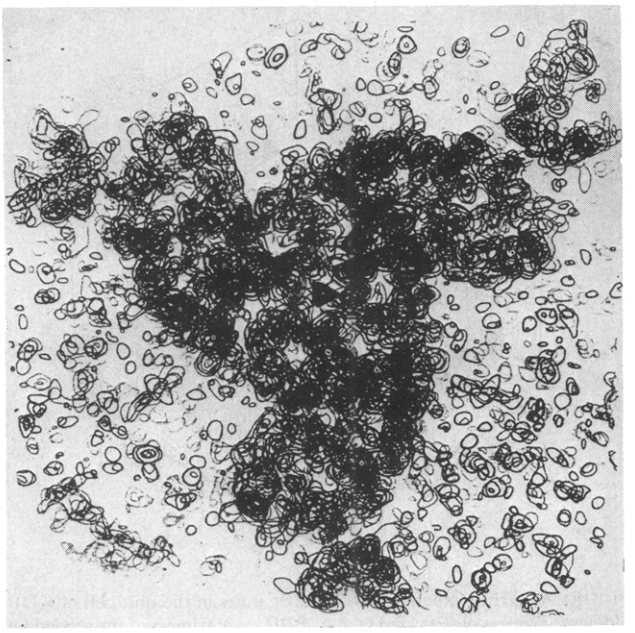


FIGURE 5: The 3.5-Å-resolution, electron-density map of KDPG aldolase viewed down threefold axis. Contours at intervals of $0.1e$ Å^{-3} ($1.1\sigma(\rho)$) beginning at $0.3e$ Å^{-3} ($3.4\sigma(\rho)$); threefold axis denoted appropriately.

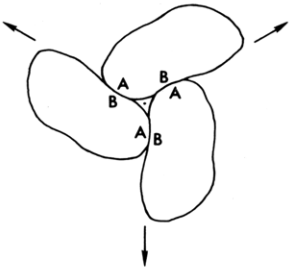


FIGURE 6: General shape and proposed manner of arrangement of subunits of the trimer of KDPG aldolase.

soid. The unique shape of KDPG aldolase trimer is in striking contrast to the more characteristic compact globular arrangements usually found in oligomeric protein systems. Conversely, the trimers of glucagon (Saski et al., 1975) and bacteriochlorophyll protein (Fenna and Matthews, 1975) display a more closed triangular shape with a central cavity. The large mother liquor content of crystals of KDPG aldolase and the characteristic shape of the trimeric molecule are similar in behavior to that observed with transfer RNA and immunoglobulin crystal and molecular structures.

The absence of disulfide bridges (Robertson et al., 1971a) and thus main chain branching facilitated the tracing of the polypeptide chain of KDPG aldolase in the electron-density map. However, in following the chain, it became inevitable that assumptions had to be made at certain points where the contours were not connected and/or more than one possibility existed as to which path to follow. Such regions were minimal and most of the backbone could be traced unambiguously. The folding of the main chain of KDPG aldolase is shown schematically in Figure 7. The regions denoted 1 and 2 in Figure 7 are where the interpretation of the electron density was not clear and the indicated folding is based more on past experience and intuition.

Nine cylindrical regions of the electron density with significantly higher density than the average and with a diameter

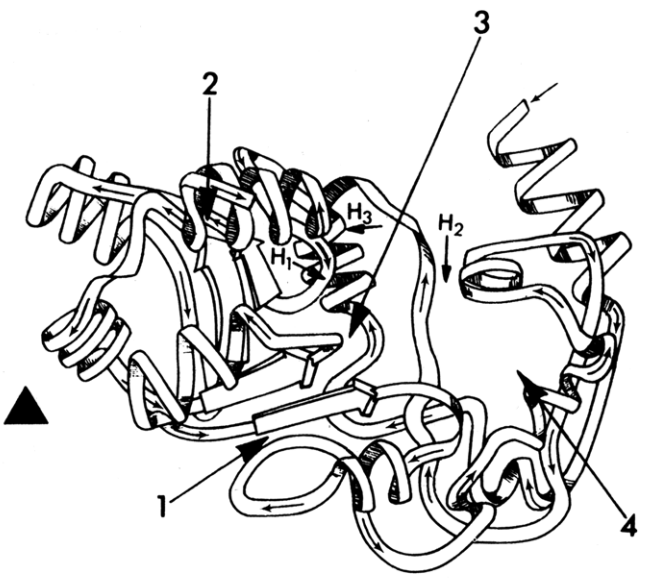


FIGURE 7: Schematic representation of the folding of a monomer of KDPG aldolase. Threefold axis indicated appropriately; H_1, H_2, and H_3 are the principal heavy atom binding sites; points 1–4 discussed in text.

of about 6 Å. have been interpreted as helical folds. These are consistent with helical regions ranging from a little over one turn to over four turns of helix. Two of the larger helical regions are found in the terminal parts of the polypeptide chain. Measuring the length of all these regions and assuming they represent α helices, the total number of residues involved in helix formation is estimated to be about 75 or about 35% of the total number of amino residues of KDPG aldolase. This is consistent with the results of optical rotary dispersion measurements (ORD) of KDPG aldolase which suggest a helical content of 31–33%. There are other regions in the map which might be helical but, since their height is not especially large, they were not included in the foregoing assessment. The relatively large helical content is consistent with the relatively large content of amino acid residues in the enzyme that are known to be good helix formers (Kotelchuck and Scheraga, 1969; Wu and Kabat, 1971, 1973) (31 Ala, 22 Leu, 20 Glu or 32% of the amino acid composition).

More recently, amino acid residues have even been classified as to their ability to form or disrupt β-sheet structures (Chou and Fasman, 1974a,b). Although there is a fair proportion of the former in KDPG aldolase (18 Ile, 15 Val, 7 Met or 18% of the composition), there is also a large proportion of β-sheet terminator (20 Glu) which would seem to prevent the formation of β structure. The latter is in agreement with the present observation that there are only two regions of the KDPG aldolase folding which look like β structures (Figure 7): these consist of a pair of two chains running parallel to each other for about 18 Å.

It was also easy to distinguish about 25 relatively long segments of electron density which obviously extend from the backbone of the molecule toward the interior of the molecule. These are most likely side chains of residues like Leu and Ile, which number 22 and 18, respectively, in the enzyme. In addition, there are about 23 segments which are probably large polar side chains of residues like Arg, Glu, and Lys because they extend toward and into the solvent region surounding the molecule. Once again, the observation is consistent with the large number of long polar side chains per subunit of KDPG aldolase (15 Arg, 20 Glu, 7 Lys). These long side chains should be useful in fitting a known 48-peptide sequence including the

active site Lys (Tsay and Wood, 1976) to the electron-density map and should be even more useful when the complete sequence is known.

A most striking feature of the folding of the molecule is the existence of a long empty channel about 9 × 9 × 30 Å which skews through the subunit near its center at about 45° to the threefold rotation axis (point 3 of Figure 7). Two parallel chains form part of the wall of this channel and run approximately parallel to its axis; however, except for this, the wall has no other obviously identifiable features. There is more than a passing interest associated with this void since three of the heavy atom binding sites of the mercury derivatives are located within it: mercury site 2 is at an entrance, site 3 further inside, and site 1 is near the midpoint of the channel (H_1, H_2, H_3 in Figure 7). Since compounds like HgS and EHgTS have been reported to bind to sulfhydryl groups and the subunit contains four cysteine residues, two readily accessible to Elman's reagent, two buried, the former might well be near heavy atom sites 2 and 3, whereas H_1 might be one of the buried cysteines.

Both ends of the central channel terminate with fairly large cavities, one about 10 × 10 × 15 Å located near the threefold axis and the other located at the other end about 8 × 10 × 25 Å (point 4 of Figure 7). Both of these voids are close to the surface in a subunit; however, only the larger void, which is a complicated arrangement of loose loops, is near the surface in the trimeric structure where it is located in the outer lobe of the subunits and can even be seen in Figure 5. Although it is not yet possible to attribute any significance to these features, it is interesting to note that the size of the larger cavity is such that it could accommodate a molecule the size of a hexose sugar which is, in fact, the substrate of the enzyme. Studies to locate the substrate binding site are underway.

(*iv*) *Quaternary Interactions.* The regions of the subunits near the threefold rotation axis which make close intersubunit contacts and represent the quaternary interactions of the trimer are characterized by a high degree of organization in the sense that most of the polypeptide chain here is folded into helicies, sharp turns, and the β structures. The extended arms or lobes of the trimer contain two of the larger helical regions (four and three turns, respectively) and the loose loop arrangement just mentioned. The quaternary interactions involve the side chains of helices and, to a lesser extent, the helical arrangements themselves.

Acknowledgment

We thank Dr. Geoffrey C. Ford, Department of Biological Sciences, Purdue University, for calculating the electron-density maps perpendicular to the threefold axes and Diane Ersfeld and Paul Kuipers for growing the crystals. Lastly, we acknowledge the many stimulating conversations and discussions we have had with Professor W. A. Wood throughout the course of this work.

References

Barran, L. R., and Wood, W. A. (1971), *J. Biol. Chem. 246,* 4028.

Blow, D. M., and Crick, F. H. C. (1959), *Acta Crystallogr. 12,* 794.

Chou, P. Y., and Fasman, G. D. (1974a), *Biochemistry 13,* 11.

Chou, P. Y., and Fasman, G. D. (1974b), *Biochemistry 13,* 222.

Colman, P. M., and Matthews, B. W. (1971), *J. Mol. Biol. 60,* 163.

Eagles, P. A. M., Johnson, L. M., Joynson, M. A. McMurray, C. H., and Gutfreund, H. (1969), *J. Mol. Biol. 45,* 533.

Fenna, R. E., and Matthews, B. W. (1975), *Nature (London) 258,* 573.

Grazi, E., Meloche, H. P., Martinez, G., Wood, W. A., and Horecker, B. L. (1963), *Biochem. Biophys. Res. Commun. 10,* 4.

Hammerstedt, R. H., Mohler, H., Decker, K. A., and Wood, W. A. (1971), *J. Biol. Chem. 246,* 2075.

Hansen, K. R. (1966), *J. Mol. Biol. 22,* 405.

Ingram, J. M., and Wood, W. A. (1966), *J. Biol. Chem. 241,* 3256.

Kartha, G., and Parthasarthy, R. (1965), *Acta Crystallogr. 18,* 749.

Kotelchuck, D., and Scheraga, H. A. (1969), *Proc. Natl. Acad. Sci. U.S.A. 62,* 14.

Lipscomb, W. N., Coppola, J. C., Hartsuck, J. A., Ludwig, M. L., Muirhead, H., Searl, J., and Steitz, T. A. (1966), *J. Mol. Biol. 19,* 423.

Mavridis, I., and Tulinsky, A. (1975), Abstracts, Tenth International Congress on Crystallography, August 7–15, 1975, Amsterdam, p 26.

Meloche, H. P. (1970), *Biochemistry 9,* 5050.

Meloche, H. P., and Glusker, J. P. (1973), *Science 181,* 350.

Meloche, H. P., Mehler, L., and Wurster, J. M. (1975), *J. Biol. Chem. 250,* 6870.

Meloche, H. P., and Wood, W. A. (1964a), *J. Biol. Chem. 239,* 3511.

Meloche, H. P., and Wood, W. A. (1964b), *J. Biol. Chem. 239,* 3515.

Mohler, J., Decker, K. A., and Wood, W. A. (1972), *Arch. Biochem. Biophys. 151,* 251.

Monod, J., Wyman, J., and Changeaux, J.-P. (1965), *J. Mol. Biol. 12,* 88.

North, A. C. T., Phillips, D. C., and Mathews, F. S. (1968), *Acta Crystallogr., Sect. A, 24,* 351.

Robertson, D. C., Altekar, W. W., and Wood, W. A. (1971b), *J. Biol. Chem. 246,* 2084.

Robertson, D. C., Hammerstedt, R. H., and Wood, W. A. (1971a), *J. Biol. Chem. 246,* 2075.

Rose, I. A., and O'Connell, E. L. (1967), *Arch. Biochem. Biophys. 118,* 758.

Sasaki, K., Dockerill, S., Adamiak, D., Tickle, I. J., and Blundell, T. L. (1975), *Nature (London) 257,* 751.

Tsay, D., and Wood, W. A. (1976), *Arch. Biochem. Biophys.* (in press).

Tulinsky, A., Mani, N. V., Morimoto, C. N., and Vandlen, R. L. (1973), *Acta Crystallogr., Sect. B 29,* 1309.

Vandlen, R. L., Ersfeld, D. L., Tulinsky, A., and Wood, W. A. (1973), *J. Biol. Chem. 248,* 2251.

Vandlen, R. L., and Tulinsky, A. (1971), *Acta Crystallogr., Sect. B. 27,* 437.

Warren, S. G., Edwards, B. F. P., Evans, D. R., Wiley, D. C., and Lipscomb, W. N. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 1117.

Wood, W. A. (1972), *Enzymes, 3rd Ed. 7,* 281.

Wyckoff, H. W., Doscher, M., Tsernoglou, D., Inagami, T., Johnson, L. N., Hardman, K. D., Allewel, N. M., Kelly, D. M., and Richards, F. M. (1967), *J. Mol. Biol. 27,* 563.

Wu, T. T., and Kabat, E. A. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 1501.

Wu, T. T., and Kabat, E. A. (1973), *J. Mol. Biol. 75,* 13.

Zeppezauer, M., Eklund, H., and Zeppezauer, E. S. (1968), *Arch. Biochem. Biophys. 126,* 564.